



Tréma

40 | 2013

L'école au défi de la mondialisation

L'évaluation de l'éducation dans le contexte de mondialisation : le cas français

Elisabeth Chatel



Édition électronique

URL : <http://journals.openedition.org/trema/3058>

DOI : 10.4000/trema.3058

ISSN : 2107-0997

Éditeur

Faculté d'Éducation de l'université de Montpellier

Édition imprimée

Date de publication : 1 décembre 2013

Pagination : 28 - 40

ISSN : 1167-315X

Référence électronique

Elisabeth Chatel, « L'évaluation de l'éducation dans le contexte de mondialisation : le cas français », *Tréma* [En ligne], 40 | 2013, mis en ligne le 01 décembre 2015, consulté le 01 mai 2019. URL : <http://journals.openedition.org/trema/3058> ; DOI : 10.4000/trema.3058

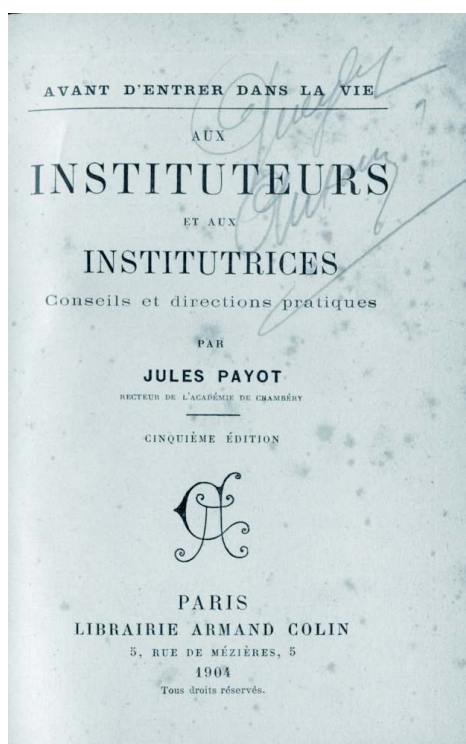
Ce document a été généré automatiquement le 1 mai 2019.

Tréma

L'évaluation de l'éducation dans le contexte de mondialisation : le cas français

Elisabeth Chatel

- 1 En éducation, comme dans beaucoup d'autres domaines de la vie sociale, l'évaluation a connu une extension extraordinaire ces vingt dernières années. Cette évolution accompagne un mouvement de mondialisation des sociétés, sans qu'on puisse, pour autant, établir de l'un à l'autre un simple lien causal. Plusieurs auteurs remarquent (Martucelli, 2012 ; Demailly, 2012), que l'évaluation porte un espoir de justice et d'efficacité. Elle entend prendre en considération des conséquences non prévues de l'action, sans renoncer à l'idéal de rationalisation ; elle vise à réintroduire une possibilité de maîtrise du monde en acceptant le fait que l'action comporte de l'incertitude et du risque. Cependant cet idéal d'une évaluation juste et efficace n'est pas aisément mis en œuvre car elle suppose une forme de collaboration des acteurs avec le processus d'évaluation de leurs actes (Demailly, 2001). Retraçant l'histoire de l'évaluation des politiques éducatives en France, X. Pons (Pons 2011) considère qu'en la matière la France reste dans un entre deux, relativement à d'autres pays, entre une optique franchement libérale tendant à utiliser l'évaluation comme outil d'une régulation quasi marchande de l'éducation, ou au



contraire comme un moyen de réflexion des responsables éducatifs dans une régulation restant centralisée et bureaucratique. Il s'interroge sur le sens d'un processus qui peut à la fois être perçu comme le vecteur d'une marchandisation de l'école et d'essor du néolibéralisme accompagnant la mondialisation des économies et sociétés pour les uns et comme l'outil d'une nouvelle gestion publique et de la réforme de l'État pour d'autres. Cet article se propose, en s'appuyant sur la littérature existante, de retracer et interroger un des aspects corrélatif au processus d'institutionnalisation de l'évaluation : le développement de la quantification des résultats des élèves.

- 2 Ce processus de quantification est introduit dans le processus éducatif avec l'évaluation car il est supposé apporter une objectivation. Or la quantification des résultats éducatifs suppose la mesurabilité et la commensurabilité des acquis de l'éducation, postulats trop peu interrogés. L'évolution que nous allons retracer, tant de la mise en place de ces quantifications que de leurs usages, montre que la réflexivité à l'égard de ces outils, parfois souhaitée par les décideurs, n'est pas au rendez-vous. Ainsi, et compte tenu d'une assez faible appropriation des outils d'évaluation par les divers acteurs de l'éducation, on peut se demander quelles sont les conséquences éducatives d'une telle évolution de l'évaluation.
- 3 Dans une première partie nous retracerons les grandes évolutions de l'évaluation de l'éducation dans le cas français en portant attention plus particulièrement aux évaluations « cognitives », et à l'extension de leur quantification. Dans une deuxième partie nous interrogerons quelques effets de ces évolutions en privilégiant le cas français.

I. L'évaluation de l'éducation vers la mesure d'un résultat

- 4 Nombreux sont les ouvrages et articles retraçant l'évolution de l'évaluation de l'éducation. Cette littérature établit le lien entre l'extension du phénomène de « gouvernement par les nombres¹ » et le contexte de mondialisation, elle montre l'influence des grandes organisations internationales telles l'OCDE ou la Banque mondiale dans ce processus (Normand, 2011). Ces organisations internationales favorisent une nouvelle régulation de l'action publique, agissant moins par la norme et plus soucieuse d'efficacité et de résultats, ceci allant de pair avec une forme de décentralisation et déconcentration administrative. Les auteurs pointent la pénétration croissante du raisonnement économique dans l'argumentaire de ces institutions.
- 5 De façon synthétique certains de ces éléments vont être repris dans ce qui suit, en insistant plus particulièrement sur la caractéristique suivante : l'alignement de l'évaluation éducative sur la mesure des acquis cognitifs des élèves, exprimés en terme de « compétences ».

I. 1 De la statistique vers l'évaluation

- 6 L'évaluation du système éducatif, en tant qu'entité globale, est relativement récente. Elle débute avec la mise en place d'un service statistique en éducation dans la tradition de la statistique française, et avec l'appui du personnel de l'INSEE. Dans les années 1970 ce service a une mission de connaissance et de prévision. Il offre des éléments descriptifs des

caractéristiques et des évolutions du système scolaire, alors que celui-ci connaît une grande extension de son accès et que la scolarité s'allonge.

- 7 Le mot évaluation prend une place centrale dans le vocabulaire administratif avec la création de la Direction de l'évaluation et de la prospective au Ministère de l'Éducation nationale en 1987. Mais auparavant diverses sortes d'évaluation des acquis des élèves ont été tentées au cours des années 1980. Ces premières évaluations standardisées des acquis des élèves permettent de mettre au point des outils, parfois en collaboration avec des praticiens, dans un bouillonnement d'initiatives issues de divers foyers (Pons, 2011).
- 8 Les évaluations de masse, ou « externes », des connaissances des élèves deviennent un outil de ces nouvelles statistiques. Par évaluation externe il faut entendre des évaluations de tous les élèves ou d'un échantillon représentatif d'entre eux à un niveau donné. Cet échantillon permet d'inférer les résultats de l'ensemble des élèves de ce niveau. Ces évaluations sont des épreuves auxquelles les élèves sont soumis, elles sont construites par des experts. C'est pourquoi elles sont dites externes, et elles sont standardisées. Elles s'institutionnalisent après 1989 sous deux formes : évaluation diagnostique et évaluation bilan.
- 9 Les évaluations « diagnostiques », après quelques atermoiements, s'établissent alors régulièrement en début d'année de CE2, de 6^e puis de seconde. Elles sont conçues pour aider les professeurs à détecter les difficultés de leurs élèves de façon à y remédier (Levasseur, 1996). Elles sont également supposées soutenir la communication des enseignants avec les parents, notamment en début d'année.
- 10 Les évaluations bilans visent une connaissance de niveau macro-social pour servir au débat et à la décision politique sur les acquis cognitifs des élèves. Cependant, elles sont irrégulières dans leur périodicité, et, malgré l'important travail méthodologique des agents de la DEP, elles ne permettent pas, à cette époque, de comparaisons chronologiques.

I. 2. De l'évaluation du système aux tentatives de pilotage par les résultats

- 11 À la fin des années 1980 l'évaluation de l'éducation change de nature en relation avec les évolutions des formes de gestion du système scolaire. La décentralisation va de pair avec un mouvement plus vaste de modernisation de l'État. L'autonomie croissante donnée aux établissements scolaires du secondaire en est l'expression. Devenus établissements publics locaux d'enseignement (EPL) en 1985, ils acquièrent une responsabilité encadrée. Ils doivent élaborer leur projet, se fixer des objectifs, sur l'atteinte desquels ils seront évalués. La DEP soutient cette démarche en élaborant de nouveaux indicateurs, les Indicateurs de Pilotage des établissements secondaires (IPES). Ces indicateurs contiennent, à côté de données descriptives (âge, redoublement, origine sociale des élèves, moyens des établissements etc.), des données relatives aux résultats cognitifs des élèves. Les données relatives aux résultats des élèves sont exprimées en termes de valeur ajoutée de l'établissement (IVAL). Les services de la DEP cherchent, en effet, à rendre plus scientifiques les indications sur les résultats que ne le font les palmarès publiés dans la presse. Ces statisticiens ayant une formation économique, ils empruntent le vocabulaire de la production pour mesurer les apports propres à chaque établissement, la valeur ajoutée. Les indicateurs de valeur ajoutée visent à éliminer l'effet du niveau initial des

élèves et celui des différences de sélection des établissements au cours du second cycle. Il est vrai que certains établissements obtiennent de bons résultats parce qu'ils sélectionnent à l'entrée ou évacuent les moins bons élèves, lorsque la concurrence entre établissements joue en leur faveur. L'indicateur cherche à capter l'apport propre de l'établissement, du point de vue de la préparation à l'examen (brevet, baccalauréat). Cependant, d'une part ces chiffres ne sont pas d'une lecture aisée, d'autre part ces tentatives de mesure d'un effet propre de l'entité éducative, que ce soit l'établissement ou même l'équipe pédagogique d'une classe, ne sont pas totalement abouties (Merle, 1999 ; Felouzis, 2005). De plus, s'inspirant du courant du capital humain en économie, ces travaux font comme si l'éducation pouvait entrer dans la métaphore de la fonction de production. Or cette posture théorique d'identification de l'éducation à une production mesurable mérite une interrogation critique (Chatel, 2001).

- 12 Dans les années 1990 l'évaluation des « résultats » des élèves se développe donc, tant au niveau global qu'à des niveaux décentralisés, comme celui de l'établissement, mais la quantification de ces résultats s'opère avec précaution. C'est le cas en ce qui concerne les indicateurs des pilotages des établissements. Quant aux données macro éducatives, elles sont assorties de commentaires méthodologiques qui alertent sur la valeur des comparaisons chronologiques, de même pour les comparaisons internationales. Celles-ci se diffusent à ce moment là sous l'égide d'une association internationale, l'IEA, qui propose des évaluations périodiques en mathématiques, en sciences et compréhension de l'écrit dans la langue maternelle. Les experts français et au ministère s'en méfient, ils soulignent les problèmes liés à la traduction et au biais culturel.
- 13 L'approche de l'évaluation qui est promue à cette époque est de type réflexif. Il s'agit d'élaborer un jugement sur l'état de l'école, ou sur tel ou tel de ses aspects. Cette façon d'envisager l'évaluation est théorisée par Claude Thelot directeur de la DEP de 1990 à 1997, par la figure du « miroir » que l'évaluation doit tendre aux acteurs de l'éducation, pour leur permettre d'analyser les résultats de leurs actions (Thelot, 2002). À côté des évaluations diagnostiques qui s'adressent principalement aux enseignants et aux parents dans les établissements et les classes, de nombreuses publications de la DEP voient les jours dans les années 1990. Elles mettent à la disposition du public un ensemble de données d'accès et de lisibilité aisés et constituent une aide à la connaissance du système, de ses évolutions, de ses résultats. Ces publications se veulent outils d'aide à la décision éducative, à tous les niveaux.

I. 3. Montée des comparaisons internationales et du thème de la performance

- 14 Dans les années 2000 la démarche évaluative prend de l'envergure dans un contexte national et international de crise économique et de concurrence exacerbée. Au niveau macro-social les gouvernements souhaitent rétablir une croissance économique et voir le chômage diminuer ; au niveau micro-social ils ont le souci de l'efficacité des dépenses publiques compte tenu de l'importance des déficits.
- 15 Or, au niveau macro-social, le développement éducatif peut, en effet, être considéré comme une source de croissance économique potentielle. La théorie économique du capital humain puis celle de la croissance endogène font des connaissances acquises et des qualifications des populations une ressource, un gisement de productivité (Gurgand,

2005). Il importe donc aux pouvoirs publics de mieux connaître et contrôler le niveau relatif d'éducation des populations.

- 16 Ils recherchent, au travers ces comparaisons, les meilleurs modèles éducatifs sur lesquels s'aligner. Cette façon de penser est confortée par l'OCDE, qui promeut l'enquête PISA (program for international student assesment) et, conjointement, par la commission européenne. Les évaluations globales des connaissances des élèves prennent de l'ampleur. Elles sont reprises et techniquement conçues pour mieux connaître l'évolution chronologique du niveau des élèves au plan national après 2003 (Évaluation CEDRE par disciplines et sur échantillon, en CM2 et fin de 3e ; Évaluation bilan de CM2, tous les dix ans). Elles sont également des outils de comparaisons internationales du fait d'une participation plus affirmée de la France aux programmes internationaux tel PISA, à partir de 2000. PISA teste tous les trois ans un échantillon représentatif de jeunes de 15 ans, sur diverses sortes de compétences cognitives.
- 17 Au niveau micro-social, la régulation par les résultats est promue dans divers domaines d'activité qui ne sont pas proprement économiques, comme l'éducation. La théorie économique de la rationalité étaye la conviction selon laquelle les incitations sont déterminantes dans la conduite des actions humaines. Cela mène les pouvoirs publics à concevoir des indicateurs pour inciter aux comportements jugés efficaces.
- 18 Ce raisonnement préside aux remaniements des évaluations dans les années 2005. La visée antérieure de diagnostic pédagogique n'est plus d'actualité, les évaluations doivent dorénavant servir d'instrument de gestion plus efficace du système éducatif en fournissant des repères mesurables pour un pilotage rapproché des équipes et des établissements. En conséquence, les évaluations diagnostiques, supposées fournir une indication localisée sur les apprentissages des élèves et soutenir la relation éducative, sont abandonnées progressivement entre 2002 (seconde), 2006 (CE2) et 2008 (sixième). La nouvelle gestion budgétaire introduite par la loi organique relative à la loi de finances (LOLF, votée en 2001 mise en place en 2006), impose une gestion des dépenses publiques par programmes, ceux-ci étant mis au service de missions. Cette nouvelle procédure suppose que face à chaque crédit alloué il y ait des indicateurs de résultats. Ainsi sont reprises en 2011-2012, à titre expérimental, des évaluations articulées aux objectifs du socle commun de connaissances et compétences². En cohérence avec cette nouvelle logique budgétaire, ces évaluations sont maintenant effectuées par la DGESCO, direction fonctionnelle de l'enseignement scolaire. Le contrôle des résultats s'éloigne de la classe pour monter à un niveau de gestion. C'est une autre façon de concevoir le rôle des indicateurs que dans la période précédente.
- 19 Ce rapide historique montre la montée des évaluations quantifiées portant sur les acquisitions cognitives des élèves comme mode d'évaluation à la fois du système et des entités. Cette logique évaluative ne se met pas en place aisément, qu'elle soit soutenue par un projet de réflexivité comme dans les années 1990, ou tirée vers la mesure de la « performance » comme dans les années 2000. Dans les deux cas, elle tranche avec une tradition antérieure d'évaluation des personnels et des entités éducatives exclusivement de type qualitatif, qui reste également en vigueur. La mesure quantifiée des acquis des élèves se superpose aux évaluations régulières des élèves en classe par des notes et appréciations. Or ces évaluations ordinaires, qu'elles soient individuelles ou collectives comme les examens (brevet des collèges, baccalauréat etc.), sont conçues par les enseignants. En cela, elles diffèrent radicalement des évaluations visant la quantification qui sont externes et standardisées.

II. Evaluation : quels effets sur l'éducation ?

- 20 Dans cette deuxième partie nous allons examiner quelques effets de cette évolution de l'évaluation vers la quantification et la mesure des acquis cognitifs. La littérature de sciences sociales à ce sujet a tout d'abord porté sur l'analyse des causes de ce phénomène. Elle a établi une relation entre le développement de l'évaluation par les résultats et de nouvelles modalités de régulation des systèmes éducatifs des pays développés. Ceux-ci seraient passés d'une régulation de type bureaucratique, qui s'opère en contrôlant la conformité des divers agents et entités aux directives venues d'en haut, à des formes accentuées d'autonomisation des entités éducatives qu'il faut donc coordonner de façon nouvelle, par leurs résultats (Broadfoot, 2000 ; Maroy, 2004 ; Nordmand, 2011).
- 21 Cependant ce mouvement général que connaissent les systèmes éducatifs des pays développés se décline différemment de l'un à l'autre (Maroy, 2006). La France, par exemple, s'éloigne de l'ancien mode de régulation sans pour autant adopter absolument le modèle de régulation par les résultats qu'on trouve dans les pays anglo-saxons, où l'évaluation a des « enjeux forts » (Pons, 2010). En effet dans ces pays (Angleterre, USA, Canada, Nouvelle-Zélande) les résultats des évaluations externes peuvent avoir des effets directs sur les personnes et les établissements : orientations, accès, redoublements des élèves, recrutement, rémunération carrière des enseignants, moyens des établissements. En France, il s'agit principalement de donner de l'information pour documenter, orienter les projets d'établissement, communiquer avec les parents, convaincre du bien fondé d'une politique etc. De sorte que l'interprétation des chiffres, des indicateurs, leur usage par les divers acteurs a beaucoup d'importance. Nous distinguerons les effets potentiels tenant aux logiques et outils à l'œuvre, des effets déjà observés qui tiennent aux interprétations et usages effectifs des décideurs politiques et des acteurs éducatifs.

II. 1 La force des outils ?

- 22 Les évaluations de masse sont-elles aussi puissantes qu'il est attendu en tant qu'outil de mise en convergence des systèmes éducatifs nationaux et comme outil de connaissance ?

II. 1. 1 Les évaluations externes comme outils de convergence des systèmes et des curricula

- 23 L'objectif de convergence est présent dans l'action des organisations internationales. L'OCDE a notamment une influence déterminante en soutenant le programme PISA (Bottani, 2004). Ce programme, dont les résultats nationaux sont publiés, a un certain impact sur l'opinion, les débats en éducation, les politiques. Par ailleurs l'intérêt de l'Union Européenne pour les questions d'éducation est réel. Cependant, l'éducation étant un domaine réservé aux compétences nationales, ce n'est qu'assez récemment qu'elle entre pleinement dans le champ des politiques communautaires. Là aussi ce sont les préoccupations économiques qui guident la démarche et son accélération. Le conseil européen de Lisbonne en 2000 exprime le projet de faire de l'Europe l'économie de la connaissance la plus compétitive. Cette « stratégie de Lisbonne » s'accompagne d'une nouvelle méthode de coordination au sein de l'union européenne, la MOC (Méthode ouverte de coordination) qui consiste à opérer des recommandations aux pays membres

d'atteindre des objectifs précis et chiffrés nommés « benchmarks » pour favoriser les convergences des meilleures pratiques. En éducation de nombreux indicateurs sont produits pour tenter d'améliorer la qualité des systèmes éducatifs, la promotion d'une culture de l'évaluation est expressément préconisée (JO de l'Union européenne 8.7.2009, Recommandation du 18 juin 2009). Ainsi un cadre de référence européen s'élabore progressivement et impose un alignement d'objectifs de référence.

- 24 On peut donc supposer que ces procédures évaluatives poussent à la convergence des structures des systèmes éducatifs et des curricula. Cette convergence attendue est, par exemple, déplorée par Rogiers (2012) qui y voit le risque que les éducateurs, pris par la force de l'outil d'évaluation qui leur est fourni, n'alignent sur lui toutes les épreuves de certification. Il voit un risque de réduction des contenus évalués et, en conséquence, de ceux qui seront enseignés. De Ketele (2010) regrette également cette homogénéisation attendue des contenus d'enseignement, qu'il voit comme porteuse de perte de diversité culturelle. Il considère que ces évaluations, parce qu'elles cherchent précisément la standardisation, contribuent à décontextualiser les contenus.
- 25 Ce point de vue ne fait pas l'unanimité. Malet et Manguez (2013) pensent que la convergence restera formelle et procédurale. Elle n'empêchera pas le maintien de la diversité car seules les compétences de base sont testées (*basic skills, key competencies*), seules donc les minima risquent de s'uniformiser, or les systèmes éducatifs sont traditionnellement édifiés en référence à leur sommet. Il faut, selon ces auteurs, moins craindre la standardisation que la dualisation inégalitaire des systèmes éducatifs du fait de leur mise en concurrence et de l'effritement de l'idéal égalitaire.
- 26 En bref, sur les effets culturels de l'outil « évaluations externes », il y a plus de suppositions que d'études empiriques qui permettraient de documenter la thèse de l'homogénéisation des contenus, de la réduction culturelle, de la convergence ou de la croissance des inégalités comme conséquences de la généralisation de ces épreuves.

II. 1. 2. Les évaluations externes comme outil de connaissance des résultats

- 27 Ces évaluations, techniquement très élaborées, permettent-elles d'avoir une connaissance plus objectivée des niveaux acquis par les jeunes d'âge scolaire ?
- 28 À ce sujet également les avis sont multiples dans la littérature. De nombreux chercheurs se servent de ces données dans leurs travaux, ils leurs accordent une valeur suffisante pour les utiliser comme preuve de leurs hypothèses. C. Baudelot et R. Establet avaient soutenu en 1989, contre les tenants de la « baisse du niveau », l'idée que le niveau des connaissances des élèves français montait à long terme (Baudelot et Establet, 1989). Ils se sont intéressés aux résultats des enquêtes PISA de 2006 (Baudelot et Establet, 2009). Ils s'arrêtent notamment sur la montée des inégalités qu'elles révèlent du fait de la croissance du pourcentage d'élèves en grande difficulté. De cet échec ils rendent responsable « l'élitisme républicain ». Cela les conduit à valoriser le modèle nordique et notamment le système de tronc commun de scolarité obligatoire qui est en vigueur en Finlande ; on sait que la Finlande reste, année après année, championne des résultats aux enquêtes PISA. Nathalie Bulle (Bulle, 2010) partage cet intérêt pour les résultats de ces enquêtes, sans pour autant aboutir aux mêmes préconisations. Elle retient le constat de la baisse des résultats des jeunes français entre 2000 et 2006. Ce constat de baisse des résultats des jeunes de 15 ans est conforté par les évaluations nationales qui montrent la baisse manifeste des résultats des élèves de CM2 entre 1987 et 2007 sur les compétences

« lire, écrire ; compter » (DEP, 2008). Cependant ces bilans alarmants ne sauraient être reliés, pour N. Bulle, à la structure de tronc commun de la scolarité obligatoire dont la réforme ne saurait remédier au problème.

- 29 Elle argumente son point de vue en pointant d'autres caractéristiques susceptibles d'expliquer le succès finlandais : repérage rapide des élèves en difficulté, accompagnement de ces élèves en petits groupes et avec des enseignants spécialisés, spécificités de la langue finnoise, confiance du public envers les enseignants.
- 30 D'autres travaux de recherche portent sur la critique des évaluations en tant qu'instruments de mesure. Gérard Bonnet, expert des évaluations internationales durant les années 1993-1995, a posé les problèmes liés aux biais culturels et aux problèmes de traduction que ces enquêtes suscitent (Normand, 2011). Selon lui, les épreuves internationales ne seraient significatives que pour les pays de culture anglo-saxonne. Par ailleurs, cet expert interroge le postulat qu'existerait un niveau de compétences inclus dans les personnes et susceptible d'être quantifié selon une échelle unique.
- 31 Mentionnons à ce sujet un travail de l'équipe ESCOL (Bautier ; Crinon ; Rayou, Rochex, 2006) d'analyse secondaire des épreuves PISA de 2000 passées par les élèves français. Cette analyse montre l'instabilité des résultats des élèves moyens qui sont la grande majorité des élèves. Elle conforte le point de vue critique de Bonnet sur le postulat qu'existerait dans les individus une compétence stable susceptible d'être un objet à mesurer.
- 32 De nombreuses autres critiques portent sur la question de la mesurabilité des acquis cognitifs, elles pointent notamment la réduction opérée par la prétention à une mesure unidimensionnelle, elles listent tout ce qui n'est pas mesuré dans ces épreuves écrites : les compétences orales, les écrits longs, la capacité à problématiser, la dimension artistique etc. Elles soulignent que ce qui n'est pas mesuré risque l'invisibilité.

II. 2. La force des usages

- 33 Le développement des évaluations a engendré celui des recherches portant sur leurs usages.

II. 2. 1. Comment les administrations et les pouvoirs publics prennent-ils en considération ces évaluations ?

- 34 Les recherches montrent une palette variée de réactions aux enquêtes et d'usage de celles-ci selon les pays (Mons, 2009). Pons (2010) analyse la réception comparée de l'enquête PISA dans trois pays européens, la France, le Portugal et l'Ecosse. Pour la France il remarque que PISA pénètre le langage politique après 2005, le message est accepté par les hommes au pouvoir, on y cherche les clés de la réussite, d'autant que la gestion des finances publiques par les résultats s'instaure avec la LOLF. Cependant dans les trois pays cette enquête a surtout servi à légitimer des politiques déjà établies. Elle permet, selon cet auteur, une rhétorique de la réforme, sans véritable appui des réformes sur les résultats des enquêtes. Felouzis (2012) souligne également les incertitudes de l'usage politique des résultats des évaluations. Il note les biais et les contradictions des travaux légitimant certaines décisions, le peu d'impact de ceux qui montrent la croissance de la proportion d'élèves en grande difficulté au primaire. Le principe de gouvernance par les résultats ne s'applique pas vraiment, mais les gouvernements s'en servent pour

responsabiliser les acteurs de terrain et pour légitimer des réformes (comme celle du primaire en 2008 selon Mons et Pons, 2013).

II. 2. 2. Enseignants, chefs d'établissements et élèves face aux évaluations

- 35 Les travaux relatifs aux pays anglo-saxons se préoccupent des effets négatifs du *testing* (évaluation externes standardisées) sur les pratiques enseignantes : travailler trop exclusivement pour préparer les tests, privilégier le suivi de certains élèves, en négliger d'autres. Certains travaux soulignent le risque que les enseignants ne se trouvent déprofessionnalisés par le fait que les tests et les exercices sont fournis par des experts. Cela engendrerait une certaine démotivation des enseignants devenus plus exécutants que concepteurs de leurs propres outils de travail (Mons, 2009).
- 36 En France, la pression des évaluations sur les enseignants est restée relativement faible tant que l'évaluation visait à remédier aux difficultés des élèves (modules, aide personnalisée). Les quelques enquêtes effectuées dans les années 1990 sur la réception des évaluations diagnostiques montrent (Déroutet et Normand 2004) que les enseignants s'en servent assez peu, ils ne se sont pas vraiment appropriés l'outil. Avec leurs organisations syndicales, ils se disent favorables à l'évaluation, mais en pratique ils pensent que c'est une procédure lourde et coûteuse qui ne leur est pas vraiment utile (Chatel, 2005).
- 37 L'effet des nouvelles modalités de régulation introduites dans les années 2000 sur la nature même des enseignements, sur leur contenu, sur les exigences, les pratiques et l'usage du temps d'enseignement commence à être étudié par les chercheurs. La pression pour une évaluation des élèves cadrée selon ces nouvelles normes s'est renforcée par les circulaires de 2009 : à l'école primaire les enfants sont évalués dès la fin de la maternelle par ce nouveau « codage » de compétences. Ils le sont en CM2, avec l'appui des évaluations nationales, au collège les professeurs doivent remplir le livret de compétences et celui-ci servira pour l'obtention de la partie contrôle continue du Brevet. Il y a de fortes injonctions des inspecteurs pour qu'ils le fassent. De la sorte les contenus acquis sont plus strictement contrôlés, et en conséquence, ceux qui sont enseignés devraient être plus strictement encadrés. Harle et Josso (2013) ont enquêté à ce sujet auprès d'enseignants, de directeurs d'école et d'inspecteurs primaires de la région nantaise. Les professeurs des écoles avec lesquels elles ont eu des entretiens sont suspicieux vis à vis de ces évaluations. Ils y voient une façon de vouloir intervenir dans leur pédagogie et de les responsabiliser sur les résultats de leurs élèves. Au demeurant, ils considèrent que ces résultats sont « déconnectés de la vie de la classe », qu'ils ne traduisent pas leur travail, que les évaluations ne les aident pas. Divers mouvements des professeurs des écoles (le refus des primes proposées à ceux qui font les évaluations en 2007, les « désobéisseurs » en 2008 contre un risque de fichage des élèves) expriment des réticences analogues à celles que cette enquête sociologique révèle.
- 38 Anne Barrère donne un écho assez proche, provenant des chefs d'établissement secondaire, enquêtés en 2004 - 2005. Elle les montre pris dans des contradictions (Barrère, 2012). Plus que les professeurs ils adhèrent à la culture de l'évaluation, ils se donnent des objectifs pour leur établissement, mais, ce travail leur prend du temps et peut les mettre en position délicate quand ils ne les atteignent pas, sans leur permettre de comprendre les raisons d'un échec relatif. Ils se trouvent ainsi en tension entre efficacité et principe civique de mixité sociale (accueillir largement tous les publics sans sélectionner, au risque

de voir baisser les résultats et partir les familles des couches moyennes et supérieures) ; tension en eux-mêmes de se voir devenir « responsable de tout ce qui arrive », sans toujours se l'expliquer.

- 39 Concernant les élèves, la question des bienfaits méfaits du *testing* a été assez documentée ces dernières années, notamment dans les pays de culture anglo-saxonne (Mons, 2009). Certains travaux montrent un effet de croissance de score donc de progrès cognitifs attribuables au *testing*, et d'autres montrent des méfaits : croissance du stress, réduction de l'éventail des compétences travaillées en classe pour réussir à l'examen, ou encore mise à l'écart des élèves les plus faibles de façon à faire réussir ceux qui ont une marge de progression. Les revues de littérature ne permettent pas d'affirmer un impact positif de ces multiples évaluations.
- 40 En France, il n'y a guère de travaux à ce sujet. Peu aussi sur les réactions des parents d'élèves aux évaluations, néanmoins A. Van Zanten observe (Van Zanten, 2001) le peu d'intérêt que les parents accordent aux informations « froides » telles qu'elles sont offertes par les évaluations standardisées.

III. Conclusion

- 41 L'évaluation externe quantifiée est entrée dans les systèmes éducatifs. Elle porte l'ambition d'en connaître objectivement les résultats par la mesure des acquis des élèves. Ce phénomène, qui se produit dans un contexte de mondialisation des économies et sociétés, contribue à une mise en visibilité mutuelle et en concurrence des systèmes éducatifs. Pour autant l'homogénéisation internationale tant des contenus enseignés que des pratiques pédagogiques ou encore des types d'organisation éducatives supposées les plus efficaces dont ce processus est potentiellement porteur ne se produit pas nécessairement. De nombreux travaux récents montrent que les divers contextes nationaux induisent une interprétation spécifique des chiffres et que les pouvoirs publics ne pratiquent pas vraiment la gouvernance par les résultats. Ils ne réagissent pas forcément quand leurs politiques se trouvent mises en difficulté par les évaluations et utilisent ces chiffres de façon assez rhétorique.
- 42 Par ailleurs, ces travaux remarquent que les évaluations ne sont jamais demandées par les acteurs de l'éducation. Elles ne les aident pas vraiment dans leur action car, le plus souvent, ils ne participent ni à leur conception, ni à leur élaboration, elles ne prennent pas vraiment en considération la nature de leur activité. Elles peuvent même les mettre en tension et être considérées par eux comme un contrôle de plus et non un soutien. La quantification se développe sans que soit suffisamment comprise et interrogée la façon dont ces chiffres sont conçus et produits au regard de ce qu'ils prétendent mesurer. Est ainsi instauré un langage puissant parce qu'il permet des comparaisons et des totalisations, mais un langage qui reste relativement étranger à l'action qu'il est supposé guider. Tous ces repères pourraient intéresser les diverses parties prenantes, à condition qu'ils se les approprient. Cependant, particulièrement en matière éducative, il est indispensable de ne pas confondre les divers niveaux d'action et de responsabilités, de distinguer éducation et gestion. Il est donc important de connaître ce que les chiffres disent et ce qu'ils masquent pour en garder un usage réflexif et, comme le dit Desrosières (Desrosières, 1993), que la discussion à propos des chiffres permette qu'advienne du nouveau.

BIBLIOGRAPHIE

- Baudelot (Christian) et Establet (Roger), *Le niveau monte*, Seuil, 1989.
- Baudelot (Christian) et Establet (Roger), *L'élitisme républicain, l'école française à l'épreuve des comparaisons internationales*, Seuil et La république des idées, 2009.
- Barrère (Anne), « Les chefs d'établissement face à l'évaluation - Tâches, dilemmes, épreuves. », *Diversité*, n° 169, p. 105-110, 2012.
- Broadfoot (Patricia) « Un nouveau mode de régulation dans un système décentralisé : l'État évaluateur », *Revue française de pédagogie*, 2000, no 130, p. 43-55.
- Bulle Nathalie., « L'imaginaire réformateur PISA et les politiques de l'école », *Revue-Skole.fr*, 18-01-2010, consulté le 26 juillet 2013.
- Bautier (Elisabeth), Crinon (Jacques), Rayou (Patrick), Rochex (Jean-Yves) « Performances en littéracie, modes de faire et univers mobilisés par les élèves : analyses secondaires de l'enquête PISA », *Revue française de pédagogie*, n°157, Octobre-novembre-décembre 2006.
- Bottani (Norberto), « Les évaluations internationales des acquis des élèves et leur impact sur les politiques d'éducation », *Politique d'éducation et de formation*, 11, 2004/2, 9-21.
- Chatel (Elisabeth), *Comment évaluer l'éducation ? Pour une théorie sociale de l'action éducative*, Delachaux et Niestlé, 2001.
- Chatel (Elisabeth), « Les dispositifs d'évaluation en éducation et l'action enseignante : le cas des évaluations de masse », p. 137-162, in Chatel E., Kirat T., Salais R., *L'action publique et ses dispositifs*, L'Harmattan, 2005.
- DEP, *Note d'information* 08. 38, 2008.
- Demailly (Lise), « Enjeux de l'évaluation et régulation des systèmes scolaires », p. 13-30, in Demailly (Lise), *Évaluer les politiques éducatives*, De Boeck, 2001.
- Demailly (Lise) « L'évaluation, entre contrôle et réflexivité », *Diversité*, n° 169, p. 51- 58, 2012.
- De Ketele (Jean-Michel), « Face visible et caché des classements internationaux », *Revue internationale d'éducation*, Sèvres, n° 54, septembre, p. 39 - 47, 2010.
- Desrosieres (Alain), *La politique des grands nombres. Histoire de la raison statistique*, La Découverte, 1993.
- Felouzis (Georges), « Performances et 'valeur ajoutée' des lycées : le marché scolaire fait des différences », *Revue française de sociologie*, 46-1, 2005, 3-36.
- Felouzis (Georges), L'évaluation des politiques éducatives- Des principes pertinents, des résultats impertinents », *Diversité*, n° 169, p. 75- 80, 2012.
- Gurgand (Marc), *Économie de l'éducation*, La Découverte, Paris, 2005.
- Harle-Giard (Isabelle) et Josso-Perrochaud (Anne), « Les évaluations nationales CE1-CM2 en France : entre enjeux institutionnels et tensions professionnels », *Spirales*, n°51, p.45-58, 2013.
- Levasseur (Jacqueline), « L'évaluation des acquis des élèves : finalités et exploitations », *Revue internationale d'éducation de Sèvres*, n°11.

- Malet (Régis) et Mangez (E.), « Présentation », *Spirales*, n° 51, p. 3-13, 2013.
- Maroy (Christian), « Régulation et évaluation des résultats des systèmes d'enseignement », *Politique d'éducation et de formation*, 11, 2004/2, 21-36.
- Maroy (Christian), *École, régulation et marché : une comparaison de six espaces scolaires locaux en Europe*, PUF, 2006.
- Martucelli, (Daniel), « Pourquoi croyons-nous à l'évaluation ? », *Diversité*, n° 169, p. 15- 21, 2012.
- Merle (Pierre), « Les indicateurs de pilotage des établissements en France », *La Revue Tocqueville*, 2, vol XX, 45-70, 1999.
- Mons (Nathalie), « Effets théoriques et réels des politiques d'évaluation standardisée », *Revue française de pédagogie*, 169, octobre-novembre-décembre, p. 99-140, 2009.
- Mons (Nathalie), Pons (Xavier), « Pourquoi le pilotage par les résultats ? », in Mons (Nathalie), Emin (Jean-Claude) Santana (Paul), *Piloter par les résultats*, CNDP, 2009.
- Mons (Nathalie), et Pons (Xavier), « Pourquoi n'y a-t-il pas de choc PISA en France ? Sociologie de la réception d'une enquête internationale dans trois pays européens (2001-2008) », *Revue française de pédagogie*, n° 182, janvier-mars 2013, p. 9-18.
- Normand (Romuald), Derouet (Jean-Louis) « Le développement d'une culture de l'évaluation à l'éducation nationale. Comment les enseignants utilisent-ils les résultats des évaluations nationales ? », *Rapport de recherche pour le Commissariat général du plan*, INRP, 2004.
- Normand (Romuald), *Gouverner la réussite scolaire ; Une arithmétique politique des inégalités*, Peter Lang, 2011.
- Pons (Xavier), « Qu'apprend-on vraiment de PISA ? Sociologie de la réception d'une enquête internationale dans trois pays européens (2001-2008) », *Revue internationale d'éducation de Sèvres*, n° 54, septembre, p. 51-59, 2010
- Pons (Xavier), *Évaluer l'action éducative. Des professionnels en concurrence*. PUF, 2010.
- Rogiers (Xavier) « L'influence des épreuves standardisées internationales sur l'évolution des dispositifs et des représentations en matière d'évaluation des acquis des élèves », *Questions vives*, Vol. 6, n° 16, <http://questionsvives.revues.org/908>, 2012
- Thelot (Claude) « Évaluer l'École », *Études*, tome 387, p. 323-334, 2002.
- Van Zanten (Agnès) « Le rôle des évaluations dans les stratégies concurrentielles des établissements et dans les choix des parents en France et en Grande-Bretagne », p. 31-46, in Demailly L., *Évaluer les politiques éducatives*, De Boeck 2001.

NOTES

1. Selon l'expression d'Alain Desrosières (Desrosières, 2003).
2. Décret du 11 juillet 2006, media.education.gouv.fr

RÉSUMÉS

Le mouvement de mondialisation des sociétés a renforcé, notamment dans les années 1990, la préoccupation relative au niveau d'éducation comparé des populations. L'évaluation de l'éducation fait une place plus grande à la mesure des acquis des élèves. Pour autant, on ne peut pas affirmer que l'efficacité du système éducatif soit améliorée, du moins dans le cas français, ni que la réflexivité des acteurs de l'éducation scolaire, d'abord visée par ces évaluations, ne soit accrue.

Globalization has reinforced, particularly in the 1990s, the concern about the level of education. Measuring student knowledge and competences becomes more and more central in educational evaluation. However, the efficiency of the educational system has not improved, at least in the French case ; the reflexivity of educational agents, expected at the beginning of the process, seems too much missing.

INDEX

Mots-clés : benchmark, évaluation externe, évaluations de masse, indicateurs de « qualité », PISA

Keywords : external evaluation, mass estimates, PISA benchmark indicators of « quality »

AUTEUR

ELISABETH CHATEL

Maitre de conférences HDR, Sociologie, IDHE (Institutions et Dynamiques Historiques de l'Economie), CNRS, ENS de Cachan